

AD-A177 000

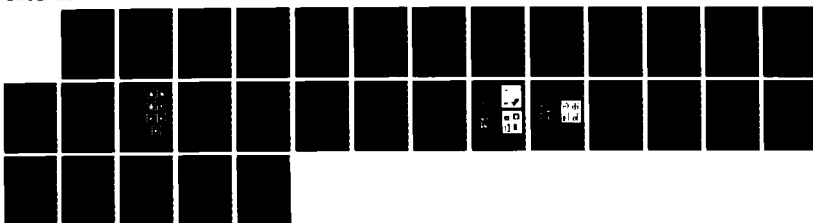
AUTOMATIC RECOGNITION AND TRACKING OF OBJECTS(U) TEXAS
UNIV AT AUSTIN COMPUTER AND VISION RESEARCH CENTER
J K AGGARWAL SEP 85 AFOSR-TR-87-0103 F49620-85-K-0007

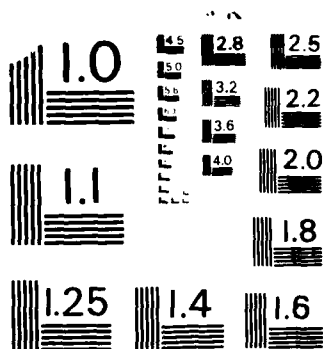
1/1

UNCLASSIFIED

F/G 17/3

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(2)

REPORT DOCUMENTATION PAGE

AD-A177 008

1b. RESTRICTIVE MARKINGS

3. DISTRIBUTION/AVAILABILITY OF REPORT

Approved for public release;
distribution unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

AFOSR-TR. 87-0103

6a. NAME OF PERFORMING ORGANIZATION

Univ. of Texas (Austin)

6b. OFFICE SYMBOL
(If applicable)

7a. NAME OF MONITORING ORGANIZATION

AFOSR

6c. ADDRESS (City, State and ZIP Code)

Austin Texas 78712

7b. ADDRESS (City, State and ZIP Code)

Bolling AFB DC 20332

BIC 410

8a. NAME OF FUNDING/SPONSORING
ORGANIZATION

AFOSR

8b. OFFICE SYMBOL
(If applicable)

NE

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

F49620-85-K-0007

8c. ADDRESS (City, State and ZIP Code)

same as 7b

10. SOURCE OF FUNDING NOS.

PROGRAM
ELEMENT NO.PROJECT
NO.TASK
NO.WORK UNIT
NO

61102F

2305

B3

11. TITLE (Include Security Classification)

Automatic Recognition and Tracking of Objects

12. PERSONAL AUTHOR(S)

Aggarwal

13a. TYPE OF REPORT
Annual

13b. TIME COVERED

FROM 1-1-85 TO 12-31-85

14. DATE OF REPORT (Yr., Mo., Day)

(9-85)

15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

Progress:

NARR. -----

PROG. - FROM U)

Jan 1-85 to Dec 31-85

18. ABSTRACT (Continue on reverse if necessary)

The Computer and Vision Research Ctr conducts a broad program of research in computer vision, image processing, and architectures for image processing. During the period of this report, several projects were completed including those on positioning and tracking of objects moving in space, parallel image processing and 3-D representation and recognition. The results on five projects are briefly presented in the report. A. Reconstruction and Matching of 3-D Objects using Quadrees/Occrees. B. 3-D Model Construction from Multiple views Using Range and Intensity Data. C. Parallel algorithms are delineated for the important task of image normalization. D. A versatile surface representation based upon the earlier volumetric description developed at our Laboratory has been formulated. E. Interpretation of Structure and Motion from Line Correspondences. During the period January 1, 1985 through December 31, 1985, our group made 20 presentations, published 9 papers in refereed journals, 11 in conference proceedings, 1 technical report and 1 non-refereed abstract. A complete listing of these activities is provided at the end of this report.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT

UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS ☐

21. ORIGINAL SECURITY CLASSIFICATION

UUUU

22a. NAME OF RESPONSIBLE INDIVIDUAL

C. LEE GILES

22b. TELEPHONE NUMBER
(Include Area Code)

767-4931

22c. OFFICE SYMBOL

NE

AFOSR-TR. 87-0103

Scientific Research Report

**AFOSR Contract
No. Contract F49620-85-K-0007**

January 1, 1985 - December 31, 1985

**Computer and Vision Research Center
The University of Texas at Austin
Austin, TX. 78712**

Principal Investigator: Professor J. K. Aggarwal

Approved for public release;
distribution unlimited.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
DISTRIBUTION STATEMENT
BY ~~TECHNICAL~~ J. KEEFER
Chief, Technical Information Division

A-1



87- 2 20 203

Scientific Report

AFOSR Contract No. F49620-85-K-0007

Abstract

The Computer and Vision Research Center conducts a broad program of research in computer vision, image processing, and architectures for image processing. During the period of this report, several projects were completed including those on positioning and tracking of objects moving in space, parallel image processing, and 3-D representation and recognition. The results on five projects are briefly presented in the report.

- A. Reconstruction and Matching of 3-D Objects using Quadtrees/Octrees,
- B. 3-D Model Construction from Multiple views Using Range and Intensity Data,
- C. Parallel algorithms are delineated for the important task of image normalization.
- D. A versatile surface representation based upon the earlier volumetric description developed at our Laboratory has been formulated,
- E. Interpretation of Structure and Motion from Line Correspondences.

During the period January 1, 1985 through December 31, 1985, our group made 20 presentations, published 9 papers in refereed journals, 11 in conference proceedings, 1 technical report and 1 non-refereed abstract. A complete listing of these activities is provided at the end of this report.

A. RECONSTRUCTION AND MATCHING OF 3-D OBJECTS USING QUADTREES/OCTREES

A.1. INTRODUCTION

The need for efficient 3-D object representations is crucial in computer vision, computer graphics, computer-aided design and other related areas. Several representation schemes have been proposed [1-4]. Representations are usually determined by the data acquisition techniques or by the type of application. For instance, a surface representation is suitable for graphic displays of opaque objects, whereas it is easier to perform operations such as matching and interference analysis with volumetric representations. A common problem with most representation techniques is that requirements for memory and processing time grow as exponential or quadratic functions of the input image size. This calls for a compact data structure that allows efficient algorithms to be implemented on it. The octree structure [5-11] with efficient tree traversal algorithms is such a candidate.

In general, an octree can be generated from a 3-D binary array using a recursive division and subdivision procedure. However, the acquisition of such a volume description is not a trivial problem. 3-D object structure can be derived from 2-D images. This task has been the primary concern of computer vision researchers. To resolve the 3-D reconstruction problem, Chien and Aggarwal [10-11] proposed a scheme to generate octrees from three orthogonal views of objects using a volume intersection technique [12-13]. Each view is extended along the associated viewing direction to form a cylinder. Each cylinder is described by a pseudo-octree. The octree of the object is generated by intersecting the three pseudo-octrees.

In this research, this algorithm is extended to generate the 'generalized octree' of an object from three known non-coplanar views, which are not necessarily orthogonal to each other. Each unit volume (voxel) associated with a node in a generalized octree is a parallelepiped, with the three sides specified by the three viewing directions.

A 'regular octree', with each voxel being a cube, is a special case of the generalized octree structure. It is known that in some cases a finite number of views is not enough to reconstruct the exact 3-D structure of an object. The more views of an object that are given, the more accurate is the description of the object that can be obtained. An object description scheme should be conducive to refinement with additional information. The proposed generalized octree structure allows subsequent refinement of the representation as additional views are available. The basic principle of the algorithm for refining octrees is similar to that of the algorithm for intersecting pseudo-octrees.

To perform object matching, the representation of the objects should be location and orientation invariant. The generalized octree structure does not meet these criteria, since it is dependent on viewing directions. A common scheme to solve this problem is to project a generalized octree onto the images planes of the three principal views (along the principal axes) to obtain the three 'principal quadrees', and to perform matching based upon the principal quadrees. Computing principal axes requires the computation of the (3×3) moment of inertia matrix comprising second order moments. To speed up processing, computation of these moments are performed based upon a 'generalized coordinate system' specified by the three viewing directions. The coordinate transformation (from the generalized coordinate system to the Cartesian coordinate system) is applied only to the moment of inertia matrix. The three principal axes can be obtained from the transformed moment of inertia matrix by computing its eigenvectors. A 'coarse' matching is performed by matching the principal quadrees of the unknown object against those of a number of models. A smaller set of models with lower degree of dissimilarities are selected. The octrees of the observed object and models are generated and a 'fine' matching is applied to octree pairs in order to identify the object. These results were presented at the Third Workshop on Computer Vision: Representation and Control, Bellaire, Michigan, October 13-16, 1985.

REFERENCES

- [1] A. A. G. Requicha, "Representations for Rigid Solids: Theory, Methods, and System," *Computing Surveys*, Vol. 12, No. 4, pp. 437-464, December 1980.
- [2] J. W. Boyse, "Interference Detection among Solids and Surfaces," *Comm. ACM*,
- [3] R. Bajcsy, *Workshop on Representation of Three-Dimensional Objects*, Univ. of Pennsylvania, May 1-2, 1979.
- [4] J. K. Aggarwal, L. S. Davis, W. N. Martin and J. W. Roach, "Survey: Representation Methods for Three-Dimensional Objects," in *Progress in Pattern Recognition*, Vol. 1, L. N. Kanal and A. Rosenfeld eds., North-Holland Publ. Co., New York, 1981, pp. 377-391.
- [5] C. L. Jackins and S. L. Tanimoto, "Octrees and Their Use in Representing Three-Dimensional Objects," *Computer Graphics and Image Processing* 14, pp. 249-270, 1980.
- [6] D. Meagher, "Geometric Modeling Using Octree Encoding," *Computer Graphics and Image Processing* 19, pp. 129-147, 1982.
- [7] I. Gargantini, "Linear Octrees for Fast Processing of Three Dimensional Objects," *Computer Graphics and Image Processing* 20, pp. 356-374, 1982.
- [8] M.-M Yau and S. N. Srihari, "A Hierarchical Data Structure for Multidimensional Digital Images," *Comm. of the ACM*, Vol. 26, No. 7, pp. 504-515, July 1983.
- [9] H. Samet, "The Quadtree and Related Hierarchical Data Structures," Technical Report MCS-83-02118, November 1983, Computer Science Department and Center for Automation Research, Univ. of Maryland, College Park, MD 20742.
- [10] W. N. Martin and J. K. Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," *IEEE Trans. on PAMI*, Vol. PAMI-5, No. 2, pp. 150-158, March 1983.
- [11] W. N. Martin, B. Gil and J. K. Aggarwal, "Volumetric Representation for Object Model Acquisition," in *Proc. NASA Symposium on Computer Aided Geometry Modeling*, NASA Conference Publication, No. 2272, April, 1983, pp. 87-94.
- [12] C. H. Chien and J. K. Aggarwal, "A Normalized Quatree Representation," *Computer Vision, Graphics and Image Processing* 26, pp. 331-346, 1984.
- [13] P. W. Likins, *Elements of Engineering Mechanics*, McGraw-Hill, New York, 1973.

B. 3-D MODEL CONSTRUCTION FROM MULTIPLE VIEWS USING RANGE AND INTENSITY DATA

B.1 Introduction

Automatic generation of computer models of the surfaces of arbitrarily shaped, three dimensional objects is an important problem in computer vision. In the past a number of different techniques have been used for representation and modeling of 3-D objects for computer vision applications [1]-[7]. However, there is an absence of a fast and robust technique for building 3-D models of arbitrarily shaped objects. In this paper, we describe a computationally efficient technique for automatic construction of 3-D models of objects given multiple views of range and intensity data.

The process of constructing 3-D models of objects involves first, integrating data or structured descriptions from multiple views of an object and then generating a representation of the complete object. In general integrating data or structured descriptions acquired from multiple views involves establishing correspondence between the views and determining the appropriate interframe transformations to register the views. The difficult and time consuming step in the above process is the matching step required to establish a correspondence. Much of the previous research efforts have been directed towards solving the difficult correspondence problem.

Several matching techniques have been developed in the past for solving this correspondence problem. Potmesil [3] generates models of 3-D objects by spatially matching 3-D surface segments describing the objects. His matching algorithm uses heuristic search to align overlapping surface segments of an object into a common 3-D coordinate system. Bhanu [1] has developed an interactive technique for constructing 3-D models of objects. The model is constructed by rotating the object through a known angle

to acquire multiple views. Coordinates of points from the multiple views are then expressed in one reference coordinate system, assuming that the interframe transformations are known a priori. Ferri and Levine [4] discuss a technique for piecing together the 3-D shape of moving objects. They construct the model of an object by first computing descriptions of visible surfaces of an object from each view of a set of multiple views. Then the interframe transformations which register these images with respect to a reference coordinate system are computed (using a feature based matching algorithm), thereby allowing for the reconstruction of surface descriptions in a world coordinate system. Boyter and Aggarwal [5] present a technique for recognizing 3-D objects using range and intensity data. They construct the model of an object by rotating the object through known angles and collecting the range and intensity line images for each object position. Bhanu et al. [2] describe a 3-D model building technique that is based on CAGD (Computer Aided Geometric Design) techniques. The 3-D data is obtained from a CAGD model of an object and the object is represented by planar approximations. The planar approximations are merged using a spatial proximity graph, to obtain a structured collection of large faces. Maggee et al. [6] present a technique for recognition of 3-D objects through intensity guided range sensing. Models are represented by a graph structure, wherein each node denotes a feature and arcs between nodes depict the geometric relationships. Boyter and Aggarwal [7] present an algorithm for recognition of polyhedra from range data. The polyhedral models are represented by 3-D coordinates of vertices, the plane equations of each face and ordered lists of vertices that bound the faces.

Most of the methods discussed above can be classified as correspondence based methods. These methods are computationally expensive due to the large search space that needs to be explored for establishing correspondence. It may be noted that none of the methods discussed above utilize information about the imaging geometry that is readily

available when constructing models. In this paper we present a technique for automatic model construction, given the range and intensity data. The technique presents a simple way of integrating information from multiple sensors namely, range and intensity. Another important feature of our method is that no point correspondences are required to determine the interframe transformations needed to express the points from each view in a common reference coordinate system.

The range and intensity data are obtained using a commercially available laser scanning system [8], which works on the principle of light sheet triangulation. The object is placed in its stable position on a flat surface called the base plane. The base plane is encoded with a pattern consisting of a single straight line. The object is positioned on the base plane such that the base plane pattern is fully or partially visible from every viewing angle. Multiple views of the object are generated by rotating the base plane about some arbitrarily fixed axis perpendicular to and on it. By observing the orientation of the base plane pattern in the intensity images of adjacent views, the interframe transformation can be easily deduced. Once the interframe transformation is known, all the (range) data are transformed into a reference coordinate system and merged. A region description of the object may then be obtained using the algorithm presented by Vemuri et al. [9]. In this representation, 3-D object surfaces are represented by regions that are a collection of surface patches homogeneous in certain intrinsic surface properties. An important aspect of this representation is that it is viewpoint independent, which is crucial for object modeling and recognition. The results were presented at the IEEE Computer Society Computer Vision and Pattern Recognition Conference at Miami Beach, Florida, 1986.

REFERENCES

- [1] B. Bhanu, "Representation and Shape Matching of 3-D Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, May 1984, Vol. PAMI-6, No. 3, pp. 340-351.
- [2] B. Bhanu, T.C. Henderson and S. Thomas, "3-D Model Building Using CAGD Techniques," *Proceedings CVPR*, 1985, San Francisco, California, pp. 234-239.
- [3] M. Potmesil, "Generating Models of 3-D Objects by Matching 3-D Surface Segments," *Proceedings 8th IJCAI*, Aug. 1983, Karlsruhe, West Germany, pp. 1089-1093.
- [4] F.P. Ferrie and M.D. Levine, "Piecing Together the 3-D Shape of Moving Objects: An Overview," *Proceedings CVPR*, 1985, San Francisco, California, pp. 574-584.
- [5] B.A. Boyter and J.K. Aggarwal, "Recognition with Range and Intensity Data," *Proceedings of the Workshop on Computer Vision : Representation and Control*, 1984, Annapolis, Maryland, pp. 112-117.
- [6] M.J. Maggee, B.A. Boyter, C.H. Chien, and J.K. Aggarwal, "Experiments in intensity Guided Range Sensing Recognition of Three-Dimensional Objects," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Nov. 1985, Vol. PAMI-7, No. 6, pp. 629-637.
- [7] B.A. Boyter, and J.K. Aggarwal, "Recognition of Polyhedra from Range Data," *IEEE Expert*, Spring 1986, Vol. 1, No. 1, pp. 47-59.
- [8] "White Scanner 100a Users Manual," Technical Arts Corporation, Seattle, Washington, 1984.
- [9] B.C. Vemuri, A. Mitiche, and J.K. Aggarwal, "3-D Object Representation from Range Data Using Intrinsic Surface Properties," to appear in *3-D Vision Systems*, T. Kanade, Editor.

C. PARALLEL IMAGE NORMALIZATION

C.1. *Introduction*

It is becoming apparent that architectures for image processing utilize two distinct types of processing elements. These form a two-level hierarchy where dedicated units will perform the high speed low-level operations and more flexible general purpose machines will perform the high-level operation [1]. In particular, the two-dimensional array structure has been shown to provide a high degree of performance for low-level operation. Furthermore, their regular structure makes them suitable for VLSI implementation [2],[3].

Image normalization is an important function frequently used in object recognition tasks [4],[5]. By "normalizing" the image of an object we refer to the process of creating a description that is invariant to the position, orientation, and size of the object in the image.

A mesh structure with one PE per pixel matches the structure of the image data and thus the normalization task essentially requires mapping each pixel to a new pixel location. Then the process is reduced to routing pixel data through the mesh.

C.2. *Processing Structure*

A four-neighbor connected mesh architecture is assumed where there is one PE per pixel. Each PE is capable of performing addition, multiplication, and comparison operations. In addition, it can maintain a FIFO queue. The queue will temporarily store routing information. The routing information or "pixel-data" for each PE, consists of three fields, pixel value, destination address, and adjacent boundary pixels.

C.3. *Parallel Normalization*

Basically, image normalization is a mapping process of each object pixel from its original position to its destination determined by the normalization parameters: translation vector, rotation angle and scale factor. The inside of objects have to be filled after the mapping when the scale factor is greater than 1 or the rotation angle is not an integer multiple of 90° . Before filling the inside, we have to reconnect the disconnected boundary. That is, the overall normalization process consists of calculation of destination, mapping, boundary reconnection, and filling.

Let $g(T, \theta, s)$ be the mapping function, where T is the translation vector, θ the rotation angle, and s the scale factor. If we use g directly when mapping, the boundary reconnection is quite complex and time consuming. In order to overcome this difficulty, we decompose g into three subfunctions: $g_1(T)$, $g_2(\theta)$, $g_3(s)$. The three processes, translation, rotation, and scaling, are therefore performed separately with g_1 , g_2 , and g_3 , respectively. Then the boundary reconnection is not necessary for the first two processes, in which the boundary is not disconnected. Moreover, the geometrical relationship between neighbors is not changed by $g_3(s)$. In other words, the direction in which one's neighbor will be found after the mapping is the same as that before the mapping. Therefore, we just have to store the information about which one(s) of 4 neighbors of a boundary pixel is boundary, for reconnection.

For each process (translation, rotation, scaling), the following procedures are executed in all PE's in parallel.

- (a) Calculation of destination address
- (b) Mapping of non-background pixels
- (c) Reconnection of boundary (scaling, $s > 1$)
- (d) Filling the inside of object (rotation, scaling)

Mapping

The basic control scheme of mapping is the "store-and-forward" mechanism. The mapping of a pixel is controlled by the repetitive application of basic flow control. We distinguish between two types of controls for regulating the flow of data between PE's. The first is the common flow control (CFC) illustrated in Figure C.1. All PE's transfer data to the same neighbors simultaneously. The second control mechanism is the discriminate flow control (DFC) where the array is partitioned into disjoint sets. Each set implements one of the forms of common flow controls. Some examples which are used in scaling and rotation are shown in Figure C.2 and C.3. A set of flow controls applied one at a time is called a *cycle* and each element of the cycle is called a *phase*. In a particular phase, if a PE contains data at the head of the queue to be transmitted in the direction specified by the phase, it transfers the data.

Boundary Reconnection

Local information about the direction of adjacent boundary pixels is available in each PE which holds a boundary pixel after mapping. If an adjacent pixel in one of those directions is not a boundary pixel, it is made a boundary pixel by changing its pixel value and transferring the information about the adjacent boundary pixels. This operation is repeated (due to a cycle of CFC's) until a boundary pixel is met. The same operation is similarly initiated, if necessary, in other directions. The disconnected boundary can be eventually reconnected.

Filling

For non-boundary and non-background pixels, any adjacent (4-connectivity) non-boundary pixel is made an object pixel. This operation is repeated (due to a cycle of CFC's) until the inside of the object has been filled.

C.4. Discussion

Figure C.4 contains simulation results for the image of an airplane. It can be seen that the proposed method works well. A small quantization error along the boundary of the normalized image is observed. We can consider some variations of

this method. Suppose that we do not have enough PE's and therefore have to assign a block of the image to a PE. Then each PE processes a block of the image assigned to it sequentially and exchanges pixel data with neighboring PE's. This method may be extended to gray level images as follows. A gray level object is segmented into several regions, each of which is uniform (same gray level). Each region may be considered as a binary object and normalized using the same algorithm proposed. The possible gaps between regions are handled by boundary reconnection and region filling algorithms.

In translation, routing pixels toward their destinations takes $N/2$ cycles at most where N is image size ($N \times N$). In rotation and scaling, the routing requires $o(N)$ cycles. Also, it takes $o(N)$ cycles to reconnect boundary or to fill the object. Therefore, the overall normalization needs $o(N)$ steps compared to $o(N^2)$ by a sequential method.

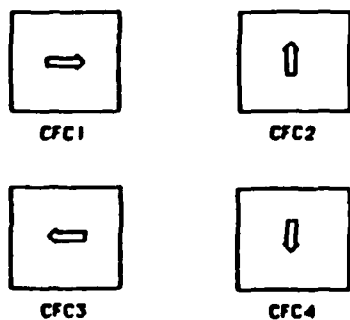


Figure C.1: Common Flow Control

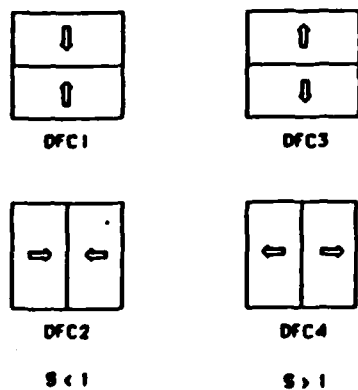


Figure C.2: Discriminate Flow Control for Scaling

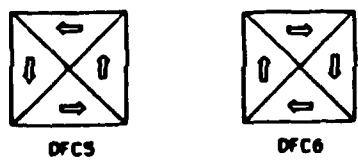


Figure C.3: Discriminate Flow Control for Rotation by Multiples of 90 degree

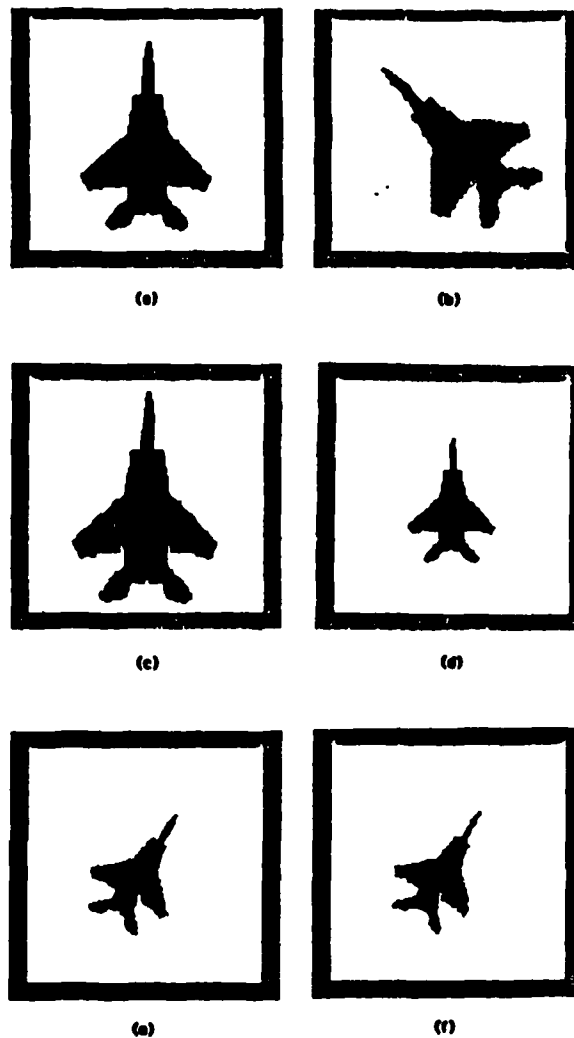


Figure C.4:

- a) Original image
- b) Rotation 45°
- c) Scaling 1.12
- d) Scaling 0.7
- e) Scaling 0.7 and Rotation -30°
- f) Rotation -30° and Scaling 0.7
- g) Rotation -45° and Scaling 0.7

C.5. References

- [1] S. Yalamanchili, K.V. Palem, L.S. Davis, A.J. Welch and J.K. Aggarwal, "Image Processing Architectures: A Taxonomy and Survey," *Progress in Pattern Recognition*, Vol. II, North Holland.
- [2] M.J.B. Duff, "Review of the CLIP Image Processing System," *Proc. of the National Computer Conference*, pp. 1055-1060, 1978.
- [3] M.J. Foster and H.T. Kung, "The Design of Special Purpose VLSI Chips," *Computer*, pp. 26-40, January 1980.
- [4] C.H. Chien and J.K. Aggarwal, "A Normalized Quadtree Representation," *CVGIP*, Vol. 26, No. 3, pp. 331-346, June 1984.
- [5] Y.C. Kim and J.K. Aggarwal, "Rectangular Coding for Binary Images," *Proc. of IEEE Computer Society Conference on CVPR*, pp. 108-113, June 1983.

D. CONSTRUCTION OF SURFACE REPRESENTATION FROM 3-D VOLUMETRIC SCENE DESCRIPTION

D.1. Introduction

This research is aimed at developing a versatile surface representation of 3-D objects from the volumetric scene description developed by Martin and Aggarwal [1], [2]. The technique we propose builds an explicit surface representation from a general description of a scene containing several occluding objects. The scene description is obtained by integrating information from several 2-D images, and is recorded as a hierarchical data structure which represents a set of planar slices of the object; each slice is characterized by a collection of 2-D shapes which define the structure at that cross section. A bottom-up approach to surface construction is adopted here. This approach involves three steps: (1) Contours on pairs of consecutive slices are examined. Contours are associated on the basis of the amount of overlapping between regions enclosed by these contours; (2) Surface elements are fit in between pairs of associated contours to establish local surface structure; and (3) These surface elements are then coalesced to form larger object facets. The resulting surface structure is recorded in a table of polygonal patches that forms the bounding surface description of the 3-D objects in the scene. Each step will be described in more detail in the following paragraphs. Some implementation results are also presented.

D.2. Contour Association

We first state the criteria for associating contours. Two contours will be associated if they are on a pair of consecutive slices, of the same sign (either both contours enclose the object regions or both contours enclose the hole regions), and the overlapping area of the two regions enclosed by the two contours is significant compared to that of the regions themselves. The 3-D scene structure is processed sequentially two slices at a time. If the overlapping area is close to that of the two regions enclosed by the two contours, both contours will be marked as processed by assigning to them a channel number to which they belong for identification. If the overlap is small compared to that of the larger of the two regions and large compared

to that of the smaller of the two regions, the larger contour can still be associated with other contours and hence is marked after all association relations are found. After the associations are completed, contours left unmarked possibly belong to different objects in the scene and are assigned new channel numbers. In this process, sequences of associated contours are recorded by attaching a unique tag (channel number). Subsequent triangulation process and surface hierarchy construction are performed over each channel independently.

D.3. *Local Surface Construction*

After the association relationships between contours on consecutive slices are identified, we will then construct the bounding surface structure between pairs of associated contours. Here we preferred planar facets to curve ones because of their representation simplicity. Finding planar surface approximation between pairs of associated contours can be formulated as a triangulation process. Briefly, the triangulation process generates a collection of triangular patches between the associated pairs of contours such that their union forms a closed bounding surface. The vertices of the triangles are the boundary points on a pair of associated contours. Triangulation of boundary points can be accomplished by constructing a graph (or matrix) representation in which the row and column indices correspond to the sequence number of the contour points on a pair of associated boundaries. A closed surface representation can then be formed by selecting a proper set of triangles such that the corresponding edges form a connected path of length $m+n$ in the graph, where m and n are the length of the previous and current contours respectively. Our method of finding such a path (or a triangulation) is based on the observation of correlation of merit assignments among neighboring triangles. This observation suggests that merit assignment should go through an iterative updating process or a 'relaxation' process to incorporate that of the neighboring triangles to ascertain the final merit assignment.

The relaxation process serves as an early screening process which speeds up the final structure construction by reducing the dimension of the selection through an early pruning of the triangles with relatively low merits. However, we might be left with many promising triangles so that more than one bounding surface structures can be built from them. A final selection of the bounding surface representation is

thus needed using a search over the restricted graph of possibilities. The well known A^* search algorithm is repeated for all pairs of associated contours to produce the local bounding surface description.

D.4. *Surface Hierarchy Generation*

Instead of incorporating triangular patches found directly as primitives for representation, we establish a surface hierarchy by coalescing triangular patches with identical orientation into polygonal facets. This is because the basic triangular patches are numerous, they may not constitute a reasonable depiction of the 3-D objects. A more reasonable depiction can be achieved if we establish the surface hierarchy by coalescing the adjacent triangular patches into polygonal facets such that the orientation of the constituent triangular patches is preserved. Data reduction is also achieved through this process.

Surface hierarchy can be established by coalescing adjacent patches in two directions: horizontal and vertical. Horizontal merging coalesces adjacent patches with same orientation within the same cross section, whereas vertical merging coalesces patches resulting from the horizontal merging across the whole scene structure. The structure obtained from this merging is shown in Figure D.1. Each polygonal facet is delineated through a pointer set which defines the bounding points for the facet at each slice thus pointer set will enable us to retrieve the detailed structure of each facet for later analysis. Information about the channel number, normal direction, and the size of the patch is also recorded.

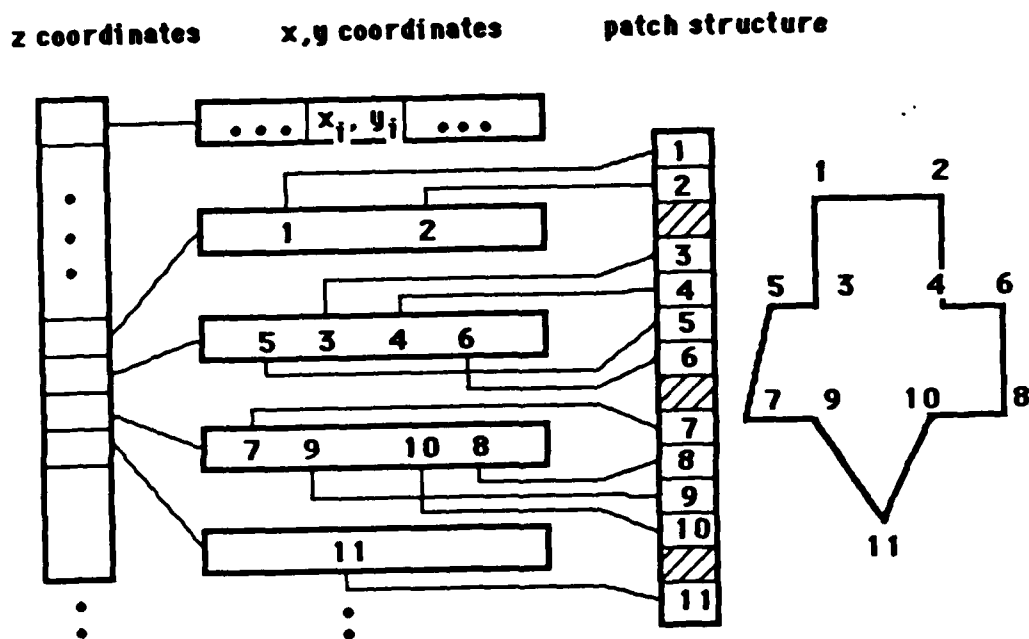


Figure D.1. The final patch structure

5. Experimental Results

Some experimental results are shown below. Figure D.2.1, D.3.1, and D.4.1 show the wire frame 3-D structure of a bus, an object with a hole, and scene with multiple objects, respectively. Figure D.2.2, D.3.2, and D.4.2 are the surface structure constructed for Figure D.2.1, D.3.1, and D.4.1, respectively, as viewed from different angles.

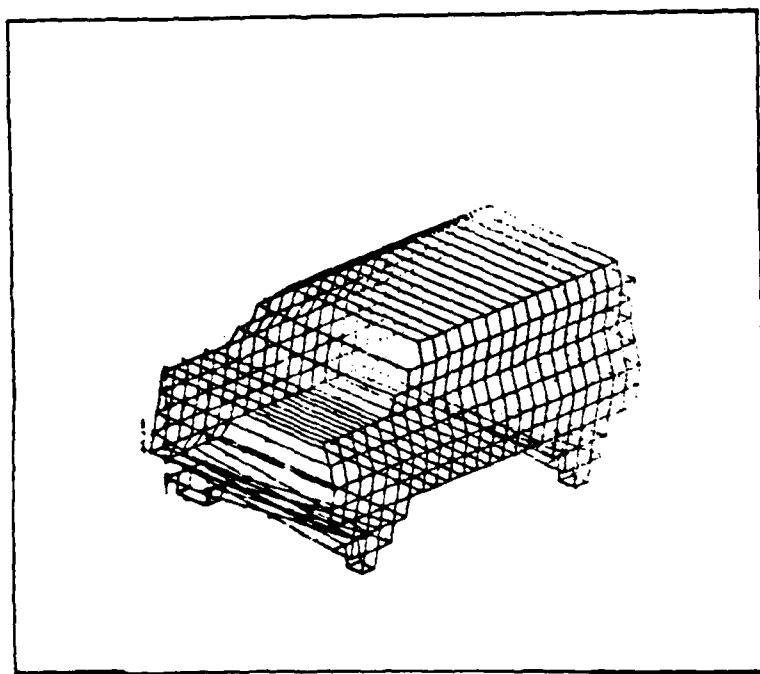


Figure D.2. (1). Wire frame 3-D structure of a bus, (2). Surface structure of the bus as viewed from the back (a), front (b), side (c), and top (d).

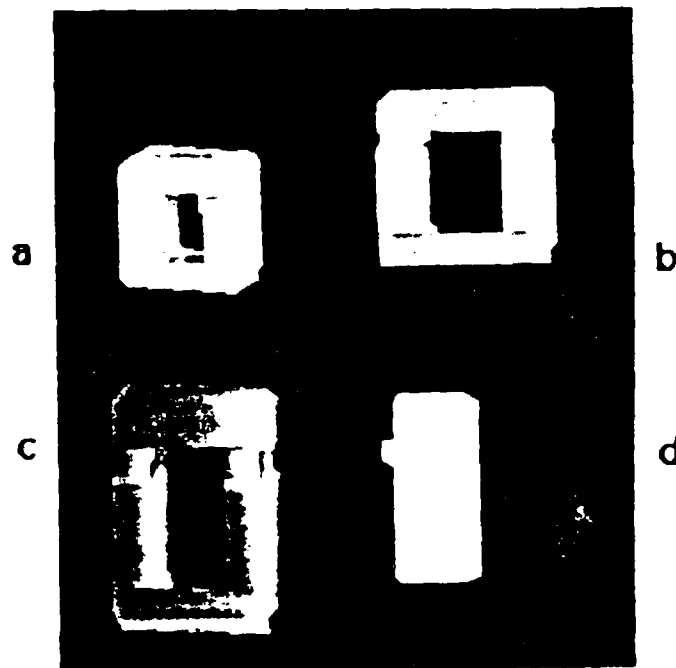
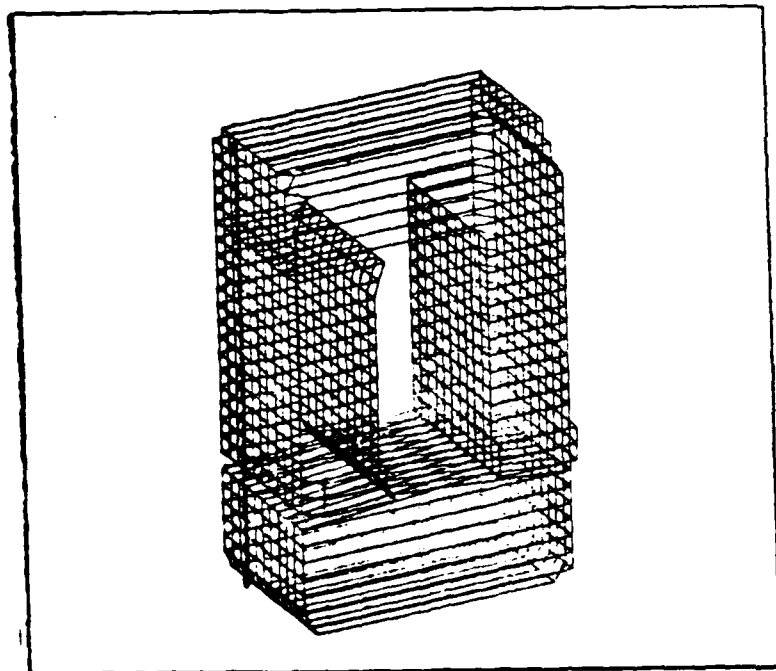


Figure D.3. (1). Wire frame 3-D structure of object with a hole, (2). Surface structure of object with a hole as viewed from side (a), (c), and (d), and front (b).

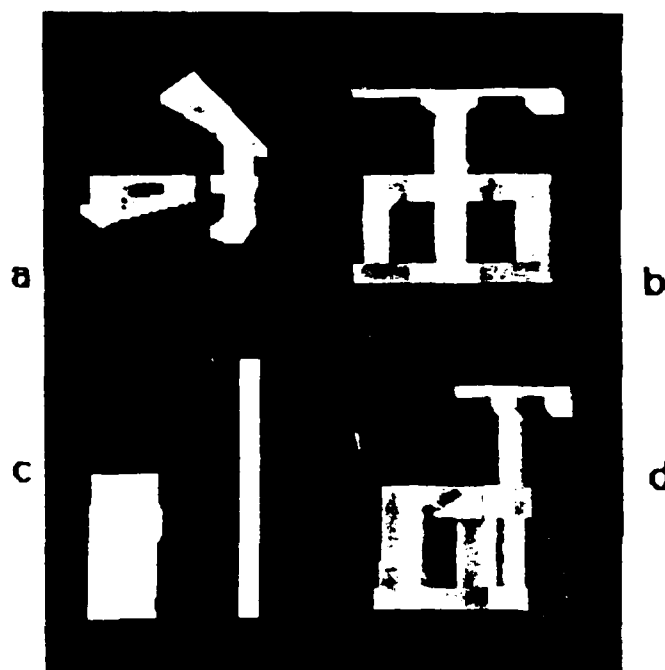
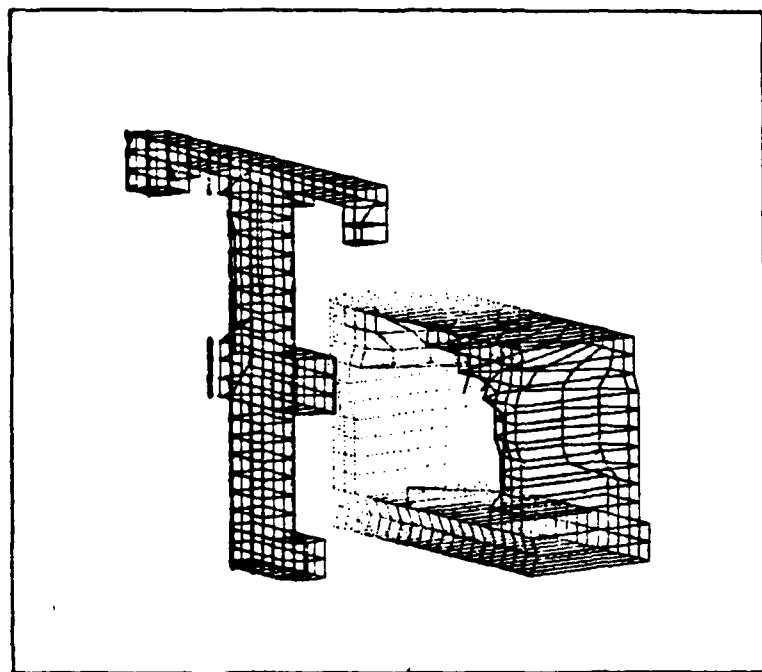


Figure D.4. (1). Wire frame 3-D structure of the scene with multiple objects, (2). Surface structure of the scene with multiple objects as viewed from side (a), (c), and (d), and front (b).

References

- [1] W.N. Martin and J.K. Aggarwal, "Occluding Contours in Dynamic Scenes," *Proc. Conf. Pattern Recog. and Image Processing*, Aug. 1981, pp. 189-192.
- [2] W.N. Martin and J.K. Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," *IEEE Trans. Pattern Recog. and Machine Intell.*, Vol. PAMI-5, No. 2, March 1983, pp. 150-158.

E. INTERPRETATION OF STRUCTURE AND MOTION FROM LINE CORRESPONDENCES

E.1. INTRODUCTION

The problem we address in this research is, in its generality, that of recovering the orientation and position of a set of lines in space from multiple views of these lines, as well as the relative displacement between the views. Research in structure and motion from images has concentrated on the use of points and optical flow. There is also a growing interest in the use of contours and range. The use of lines [1]-[3] has been relatively neglected although these may often be easier to extract from images. In [1] a method has been proposed which relies on the projective configuration of six lines in three views. In this paper we describe a method based on the observation of four lines in three views (the case of two views of any number of lines is inherently ambiguous). This method exploits the **principle of invariance of angular configuration with respect to rigid motion** in addition to the usual projective constraints. The method first solves for the orientation of lines in space. The rotational component of motion between the viewing systems is then readily recovered from these orientations. Finally, the translation components of motion (and therefore the position of lines in space) can be recovered. The results will be presented at the International Conference on Pattern Recognition in Paris in October 1986.

REFERENCES

- [1] Y. Liu and T.S. Huang, "Estimation of Rigid Body Motion Using Straight Line Correspondences," *Proc. IEEE Workshop on Motion: Representation and Analysis*, Charleston, S.C., pp. 47-51, May 1986.
- [2] A. Mitiche, S. Seida and J.K. Aggarwal, "Line-Based Computation of Structure and Motion Using Angular Invariance," *Proc. IEEE Workshop on Motion: Representation and Analysis*, Charleston, S.C., pp. 175-180, May 1986.
- [3] J.K. Aggarwal and A. Mitiche, "Structure and Motion from Images," *DARPA Image Understanding Workshop*, Miami Beach, FL, pp. 89-97, December 1985.

PRESENTATIONS, PROCEEDINGS,
AND PUBLICATIONS

Presentations, Proceedings, and Publications

Presentations

- [1] Kim, Y. C. and Aggarwal, J. K., "Finding Range from Stereo Images," presented at the *Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985.
- [2] Kim, Y. C. and Aggarwal, J. K., "Rectangular Parallelepiped Coding for Solid Modeling," presented at the *6th IASTED Int. Symp. Robotics and Automation*, Santa Barbara, California, May 1985.
- [3] Lee, S. Y., Yalamanchili, S., and Aggarwal, J. K., "Parallel Image Normalization," presented at the *Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985.
- [4] Mitiche, A., Seida, S., and Aggarwal, J. K., "Determining Position and Displacement in Space from Images," presented at the *Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985.
- [5] Wang, Y. F. and Aggarwal, J. K., "Construction of Surface Representation from 3-D Volumetric Scene Description," presented at the *Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985.
- [6] Tsukune, H. and Aggarwal, J. K., "Analyzing Orthographic Projection of Multiple 3-D Velocity Vector Fields in Optical Flow," presented at the *IEEE Computer Society Computer Vision and Pattern Recognition Conference*, San Francisco, CA, June 1985.
- [7] Vemuri, B., Mitiche, A., and Aggarwal, J. K., "3-D Object Representation Utilizing Intrinsic Surface Properties," presented at the *SIAM Conference on Geometric Modelling and Robotics*, Albany, NY, July 1985.
- [8] Aggarwal, J. K., "Computer Vision: Present and Future," presented at the *University of Illinois*, Urbana/Champaign, IL, May 1985.
- [9] Aggarwal, J. K., "Computer Vision: Present and Future," at the *University of Utah*, Salt Lake City, UT, March 1985.
- [10] Aggarwal, J. K., "Computer Vision: Present and Future," presented at the *University of Michigan and Industrial Technology Institute*, Ann Arbor, MI, February 1985.
- [11] Aggarwal, J. K., "Computer Vision: Present and Future," presented at the *GM Research Laboratories*, Warren, MI, February 1985.
- [12] Aggarwal, J. K., "Structure and Motion from Images," presented at the *Image Understanding Workshop*, Miami, Florida, December 9-10, 1985.

- [13] Aggarwal, J. K., "Structure and Motion from Images: Fact and Fiction," presented at the *IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985.
- [14] Aggarwal, J. K., "Computer Vision Research at The University of Texas at Austin," presented at the *IBM Scientific Center*, Rome, Italy, December 1985.
- [15] Aggarwal, J. K., "3-D Structure and Motion from 2-D Images," presented at the *International Conference on Advances in Image Processing and Pattern Recognition*, Pisa, Italy, December 11, 1985.
- [16] Aggarwal, J. K., "Computer Vision Research at The University of Texas at Austin," presented at *Texas Instruments*, Dallas, Texas, November 1985.
- [17] Lee, S. Y. and Aggarwal, J. K., "A Problem-Driven Approach to Parallel Image Processing: System Design and Scheduling," presented at the *1985 International Conference on Systems, Man and Cybernetics*, Tucson, Arizona, November 12-15, 1985.
- [18] Martin, W. N. and Aggarwal, J. K., "Texture Analysis and Multi-resolution Imagery," presented at the *Fourteenth Workshop on Applied Imagery*, Pattern Recognition, Baltimore, Maryland, October 1985.
- [19] Chien, C. H. and Aggarwal, J. K., "Reconstruction and Matching of 3-D Objects Using Quadtrees/Octrees," presented at the *IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985.
- [20] Wang, Y. F., Mitiche, A., and Aggarwal, J. K., "Inferring Local Surface Orientation with the Aid of Grid Coding," presented at the *IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985.

Papers

- [21] Yalamanchili, S. and Aggarwal, J. K., "Analysis of a Model for Parallel Image Processing," *Pattern Recognition*, Vol. 18, No. 1, 1985, pp. 1-16.
- [22] Yalamanchili, S. and Aggarwal, J. K., "A System Organization for Parallel Image Processing," *Pattern Recognition*, Vol. 18, No. 1, 1985, pp. 17-29.
- [23] Mitiche, A. and Aggarwal, J. K., "Image Segmentation by Conventional and Information Integrating Techniques: A Synopsis," *Image and Vision Computing*, 1985.
- [24] Yalamanchili, S. and Aggarwal, J. K., *Proceedings of the Workshop on Algorithm Guided Parallel Architecture for Automatic Target Recognition*, Leesburg, Virginia, July 16-18, 1984, prepared by C. L. Giles and A. Rosenfeld, February, 1985.

- [25] Yalamanchili, S., Palem, K. V., Davis, L. S., Welch A. J., and Aggarwal, J. K. "Image Processing Architectures: A Taxonomy and Survey," in *Progress in Pattern Recognition*, Vol. 2, Edited by L. Kanal and A. Rosenfeld, Published by North-Holland, pp. 1-37, 1985.
- [26] Magee, M. and Aggarwal, J. K., "Using Multisensory Images to Derive the Structure of Three-Dimensional Objects - A Review," *Computer Vision, Graphics and Image Processing*, Vol. 32, pp. 145-157, 1985.
- [27] Nandhakumar, N., and Aggarwal, J. K., "The Artificial Intelligence Approach to Pattern Recognition - A Perspective and an Overview," *Pattern Recognition*, Vol. 18, No. 6, pp. 383-389, 1985.
- [28] Yalamanchili, S., and Aggarwal, J. K., "Reconfiguration Strategies for Parallel Architectures," *IEEE Computer Society COMPUTER*, Vol. 18, No. 12, pp. 44-61, December 1985.
- [29] Magee, M., Boyter, B., Chien, C. H., and Aggarwal, J. K., "Experiments in Intensity Guided Range Sensing Recognition of Three-Dimensional Objects," *Pattern Analysis Machine Intelligence Transactions*, Vol. PAMI-7, No. 6, pp. 629-637, November 1985.

Conference Proceedings

- [30] Kim, Y. C. and Aggarwal, J. K., "Finding Range from Stereo Images," *Proceedings of Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985, pp. 289-294.
- [31] Lee, S. Y., Yalamanchili, S., and Aggarwal, J. K., "Parallel Image Normalization," *Proceedings of Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985, pp. 614-619.
- [32] Mitiche, A., Seida, S., and Aggarwal, J. K., "Determining Position and Displacement in Space from Images," *Proceedings of Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985, pp. 504-509.
- [33] Wang, Y. F. and Aggarwal, J. K., "Construction of Surface Representation from 3-D Volumetric Scene Description," *Proceedings of Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985, pp. 130-135.
- [34] Tsukune, H. and Aggarwal, J. K., "Analyzing Orthographics Projection of Multiple 3-D Velocity Vector Fields in Optical Flow," *Proceedings of the IEEE Computer Society Computer Vision and Pattern Recognition Conference*, San Francisco, California, June 1985, pp. 510-517.
- [35] Kim, Y. C. and Aggarwal, J. K., "Rectangular Parallelepiped Coding for Solid Modeling," *Proceedings of the 6th IASTED International Symposium on Robotics and Automation*, Santa Barbara, California, May 1985, pp. 1-5.

- [36] Mitiche, A. and Aggarwal, J. K., "Structure and Motion from Images: Fact and Fiction," *Proceedings of the IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985, pp. 127-128.
- [37] Lee, S. Y. and Aggarwal, J. K., "A Problem-Driven Approach to Parallel Image Processing: System Design and Scheduling," *Proceedings of the 1985 International Conference on Systems, Man and Cybernetics*, Tucson, Arizona, November 12-15, 1985, pp. 680-686.
- [38] Chien, C. H. and Aggarwal, J. K., "Reconstruction and Matching of 3-D Objects Using Quadtrees/Octrees," *Proceedings of the IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985, pp. 49-54.
- [39] Wang, Y. F., Mitiche, A., and Aggarwal, J. K., "Inferring Local Surface Orientation with the Aid of Grid Coding," *Proceedings of the IEEE Computer Society Third Workshop on Computer Vision: Representation and Control*, Bellaire, Michigan, October 13-16, 1985, pp. 96-104.
- [40] Mitiche, A. and Aggarwal, J. K., "Structure and Motion from Images," *Proceedings: DARPA Image Understanding Workshop*, Miami, Florida, December 9-10, 1985, pp. 89-97.

Technical Reports and Other Publications

- [41] Yalamanchili, S. and Aggarwal, J. K., "Parallel Image Processing: Methodology, Modeling and Analysis," No. TR-85-1-25, January 1985.
- [42] Aggarwal, J. K., "3-D Structure and Motion from 2-D Images," Abstracts for International Conference on Advances in Image Processing and Pattern Recognition, Pisa, Italy, December 10-12, 1985.

END

3-87

Dtic